

# REPELLERS FOR NON-UNIFORMLY EXPANDING MAPS WITH SINGULAR OR CRITICAL POINTS

KATRIN GELFERT

**ABSTRACT.** Given an ergodic measure with positive entropy and only positive Lyapunov exponents, its dynamical quantifiers can be approximated by means of quantifiers of some family of uniformly expanding repellers. Here non-uniformly expanding maps are studied that are  $C^{1+\beta}$  smooth outside a set of possibly critical or singular points.

## 1. INTRODUCTION

By a well-established technique, a  $C^{1+\beta}$  diffeomorphism that preserves a hyperbolic ergodic measure of positive entropy can be approximated gradually by compact invariant locally maximal hyperbolic sets – horseshoes. Here approximation is to be understood in terms of dynamical quantities such as the topological entropy, the topological pressure of a continuous function, Lyapunov exponents and averages of continuous functions with respect to ergodic measure that are supported on the horseshoes. In this paper we prove an analogous version in the case of a map that possesses some ergodic measure with positive entropy and only positive Lyapunov exponents and we show a gradual approximation by uniformly expanding repellers. We are interested in a quite general class of maps that are  $C^{1+\beta}$  smooth outside some set  $\mathcal{S}$  that can contain critical and singular points of a certain type or points where  $f$  is discontinuous.

Let  $f: M \rightarrow M$  be a map on a compact  $n$ -dimensional Riemannian manifold  $M$ . Let  $\mathcal{S} \subset M$  be a set that may be thought of the set of points  $x \in M$  where  $df(x)$  is either not defined or where  $df(x)$  is not invertible. Assume that  $f: M \setminus \mathcal{S} \rightarrow f(M \setminus \mathcal{S})$  be a  $C^{1+\beta}$  map. We require that  $f$  satisfies additional conditions  $(C_1)$  and  $(C_2)$  that will be specified below. Recall that  $R$  is a *uniformly expanding repeller* with respect to  $f$  if  $R$  is a compact  $f$ -invariant isolated set such that  $f|_R$  is uniformly expanding and topologically transitive. Recall that  $f|_R$  is said to be *uniformly expanding* if there exist  $c > 0$  and  $\lambda > 1$  such that for every  $n \geq 1$  and every  $x \in R$

---

2000 *Mathematics Subject Classification.* 37D40, 37D50, 37C35.

*Key words and phrases.* Pesin theory, non-uniformly hyperbolic dynamics, horseshoes, entropy.

I am grateful to Feliks Przytycki and Vilton Pinheiro for sharing their insights and acknowledge the hospitality of IMPA, Rio de Janeiro, where part of this research was done. I was supported by EU FP6 ToK SPADE2 and by the Alexander von Humboldt Foundation.

we have  $|(f^n)'| \geq c\lambda^n$ . Recall that  $R$  is said to be *isolated* if there exists an open neighborhood  $U \subset M$  of  $R$  such that  $f^n(x) \in U$  for every  $n \geq 0$  implies  $x \in R$ .

The following is the first main result of this paper.

**Theorem 1.** *Let  $f: M \setminus \mathcal{S} \rightarrow f(M \setminus \mathcal{S})$  be a  $C^{1+\beta}$  map and  $\mu$  be an  $f$ -invariant ergodic Borel probability measure satisfying  $(C_1)$  and  $(C_2)$ . Assume that  $\mu$  has positive entropy and that it admits only positive finite Lyapunov exponents that are bounded from below by some number  $\chi(\mu) > 0$ . Let  $\varphi = \varphi_1, \dots, \varphi_K: M \rightarrow \mathbb{R}$  be continuous functions.*

*For every  $\varepsilon > 0$  there exists a compact  $f$ -invariant set  $Q_\varepsilon \subset M \setminus \mathcal{S}$  such that  $f|_{Q_\varepsilon}$  is uniformly expanding and satisfies*

- (a)  $h_{\text{top}}(f|_{Q_\varepsilon}) \geq h_\mu(f) - \varepsilon$ ,
- (b) for every  $j = 1, \dots, K$

$$P_{\text{top}}(f|_{Q_\varepsilon}, \varphi_j) \geq h_\mu(f) + \int \varphi_j d\mu - \varepsilon,$$

- (c) for every  $x \in Q_\varepsilon$  and every  $j = 1, \dots, K$

$$\limsup_{n \rightarrow \infty} \left| \frac{1}{n} (\varphi_j(x)) + \varphi_j(f(x)) + \dots + \varphi_j(f^{n-1}(x)) - \int \varphi_j d\mu \right| < \varepsilon,$$

- (d) *Lyapunov exponents of periodic points satisfy*

$$\inf \{ \lambda(x, v) : x = f^{n(x)}(x) \in Q_\varepsilon \text{ and } v \in T_x M \setminus \{0\} \} \geq \chi(\mu) - \varepsilon.$$

Moreover, there exists  $m \geq 1$  such that  $f^m|_{Q_\varepsilon}$  is a uniformly expanding repeller.

Results of this type are widely referred to Katok [9] or Katok and Mendoza, see [10]. An earlier related statement for continuous and for piecewise monotone maps of the interval goes back to Misiurewicz and Szlenk [17]. Corresponding properties of holomorphic maps are shown in [21]. The case of  $C^{1+\beta}$  maps is covered in [7], see also [5] for a sketch. A related setting of dyadic diophantine approximations is established in [19]. Following similar ideas, Mendoza [15] and Sánchez-Salas [23] investigate how hyperbolic SRB-measures can be approximated by ergodic measures that are supported on horseshoes of arbitrarily large unstable dimension. Similar results in the case of holomorphic functions are derived by Przytycki in [20] and in [21, Chapter 11].

We now formulate and discuss our assumptions  $(C_1)$  and  $(C_2)$ . First, we assume that  $f$  preserves an  $f$ -invariant ergodic Borel probability measure  $\mu$  satisfying

$$(C_1) \quad \log^+ \|df\| \in L^1(\mu), \log^+ \|(df)^{-1}\| \in L^1(\mu).$$

Throughout we use the notation  $\log^+ a = \max\{\log a, 0\}$ . We want to include into our analysis maps that are Hölder continuously differentiable outside  $\mathcal{S}$ , but may have unbounded derivatives. In the presence of such singularities (uniform) Hölderness of the derivative may be lost, and similar arguments

apply to local inverses of the function  $f$ . To have some control on the Hölderness close to singular points, we require the following hypothesis to be satisfied:

(C<sub>2</sub>) There are functions  $G, H$  such that  $\log G, \log H \in L^1(\mu)$  and that for every  $x, y \in M \setminus \mathcal{S}$  with  $d(f(x), f(y)) < G(f(x))$  and  $v \in T_{f(x)}M$  we have

$$\|df(x)^{-1}(v) - df(y)^{-1}(v)\| \leq H(f(x)) d(f(x), f(y))^\beta \|v\|. \quad (1)$$

Moreover

$$\log d(\cdot, \mathcal{S}) \in L^1(\mu). \quad (2)$$

Here  $d(x, \mathcal{S})$  denotes the Riemannian distance of a point  $x$  from  $\mathcal{S}$ . The latter condition (2) is required to have under these general requirements control on the asymptotic behavior of orbits that eventually approach singularities or critical points. In fact, as it can be seen below, it is sufficient to require that  $f$  has slow return to critical points in the sense of Lemma 2.

We phrase the general condition (C<sub>2</sub>), in order to put it into the context of other commonly used approaches and mainly follow an approach in [11]. The main difference is that in [11] they assume the map  $f$  to be  $C^2$  diffeomorphism from  $M \setminus \mathcal{S}$  onto its image  $f(M \setminus \mathcal{S})$ , and they assume some control on the second derivative of  $f$ . In our setting  $f$  need not be invertible, and we require  $f$  to be  $C^{1+\beta}$  smooth outside  $\mathcal{S}$ . Similar approaches abstracting from one-dimensional maps [8] and maps with singularities [11] to higher-dimensional maps require that  $f$  is a  $C^{1+\beta}$  local diffeomorphism outside a set  $\mathcal{S}$  and is *non-flat*, that is, behaves like a power of the distance close to the singular set  $\mathcal{S}$  and in addition shows slow recurrence to  $\mathcal{S}$ . In comparison to that we will not require any more particular properties of  $f$  close to  $\mathcal{S}$ . In particular, Theorem 1 is also applicable to  $C^{1+\beta}$  maps with flat critical points.

Notice that for (1) it is sufficient to assume that  $df$  is Hölder continuous with some control on the Hölder constant. Then we can use a special version of the inverse mapping theorem for maps with Hölder continuous derivatives (see for example [4, Lemma 4.1.3]) to verify an inverse branch  $y \mapsto f^{-1}(y)$  to be of class  $C^{1+\beta}$ .

Let us now discuss some special cases and examples that fit into our settings. First let us consider a particular case of a map with non-flat critical/singular points that show a certain non-degeneracy as for example in [3].

**Theorem 2.** *Let  $f: M \setminus \mathcal{S} \rightarrow f(M \setminus \mathcal{S})$  be a  $C^2$  local diffeomorphism and assume that there are constants  $H > 1$  and  $\alpha > 0$  such that for  $x \in M \setminus \mathcal{S}$  and every  $v \in T_x M \setminus \{0\}$  we have*

$$(F_1) \quad H^{-1}d(x, \mathcal{S})^\alpha \leq \frac{\|df(x)(v)\|}{\|v\|} \leq Hd(x, \mathcal{S})^{-\alpha}$$

*and that for every  $x, y \in M \setminus \mathcal{S}$  with  $d(x, y) < d(x, \mathcal{S})/2$  we have*

(F<sub>2</sub>)

$$|\log \|df(x)^{-1}\| - \log \|df(y)^{-1}\|| \leq Hd(x, \mathcal{S})^{-\alpha} d(x, y)$$

$$|\log |\det df(x)| - \log |\det df(y)|| \leq Hd(x, \mathcal{S})^{-\alpha} d(x, y)$$

Then for any  $f$ -invariant ergodic Borel probability measure  $\mu$  that has positive entropy, admits  $\mu$ -almost everywhere only positive finite Lyapunov exponents, and satisfies  $\log d(\cdot, \mathcal{S}) \in L^1(\mu)$  the conclusions (a)–(d) of Theorem 1 are true.

*Proof.* Observe that (F<sub>1</sub>) implies that

$$|\log \|df(x)\||, |\log \|(df)^{-1}(x)\|| \leq \text{Const} + \beta |\log d(x, \mathcal{S})|$$

for all  $x$  sufficiently close to  $\mathcal{S}$ . Now from  $\log^+ \|df\| < |\log \|df\||$  we obtain  $\log^+ \|df\| \in L^1(\mu)$  and analogously  $\log^+ \|(df)^{-1}\| \in L^1(\mu)$ . Thus (C<sub>1</sub>) and (C<sub>2</sub>) are satisfied and Theorem 1 applies.  $\square$

Note that under the hypothesis of Theorem 2 Proposition 4.1 in [2] implies  $\log d(\cdot, \mathcal{S}) \in L^1(m)$  in the case that  $\mathcal{S} \subset M$  is a compact submanifold of dimension  $< \dim M$  and  $m$  the Lebesgue measure. Moreover, if  $\mu$  is absolutely continuous with respect to the Lebesgue measure and has a density  $\in L^q(m)$  for some  $q > 1$  then  $\log d(\cdot, \mathcal{S}) \in L^1(\mu)$  [2, Corollary 4.2].

**Example 1** (Cusp maps). Consider an interval  $I \subset \mathbb{R}$ , a set  $\mathcal{S} = \{s_n\}_n \subset I$ , and a map  $f: I \setminus \mathcal{S} \rightarrow I$  such that there exist constants  $\beta > 0$  and  $H > 0$  such that  $\inf_{I \setminus \mathcal{S}} |f'| > 0$  and for each connected component  $J \subset I \setminus \mathcal{S}$  for every  $x, y \in J$

$$|f'(x)^{-1} - f'(y)^{-1}| < H|x - y|^\beta.$$

Particular examples are the map  $f: [-1, 1] \setminus \{0\} \rightarrow [-1, 1]$  given by

$$f(x) = \begin{cases} 2\sqrt{x} - 1 & \text{if } x > 0, \\ 1 - 2\sqrt{|x|} & \text{if } x < 0 \end{cases}$$

as well as the Gauß map. For both maps it can be shown that the Lebesgue measure is invariant and satisfies (C<sub>1</sub>) and (C<sub>2</sub>). Certain Lorenz-like maps may also provide examples (also with criticalities and singularities such as in [14] (compare expansion estimates in [14, Section 3]).

Let us now consider the case that  $f: M \rightarrow M$  is a  $C^{1+\beta}$  endomorphism, that is,  $\mathcal{S}$  contains at most critical points of  $f$ .

**Theorem 3.** *Let  $f: M \rightarrow M$  be a  $C^{1+\beta}$  map. Then for any  $f$ -invariant ergodic Borel probability measure  $\mu$  that has positive entropy and admits only positive finite Lyapunov exponents the conclusions (a)–(d) of Theorem 1 are true.*

*Proof.* Consider an invariant ergodic probability measure  $\mu$ . Since  $f$  is  $C^1$  and hence  $\|df\|$  is bounded we obtain  $\log^+ \|df\| \in L^1(\mu)$  and we can apply the multiplicative ergodic theorem [22, Theorem 1.6]. Note that the set  $\mathcal{S}$  in this case contains only critical points of  $f$ .

Let us assume that  $\mu$  has finite Lyapunov exponents  $\lambda_1(\mu) \geq \dots \geq \lambda_{\dim M}(\mu) > 0$ . Observe that

$$m(df^n(x))^{\dim M} \leq |\det df^n(x)| \leq \|df^n(x)\|^{\dim M-1} m(df^n(x)) \quad (3)$$

for every  $x$  and every  $n$ , where  $m(df(x)) = \|(df)^{-1}(x)\|^{-1}$  whenever  $x \notin \mathcal{S}$  and  $= 0$  otherwise. The multiplicative ergodic theorem and the Birkhoff ergodic theorem (applied to  $\log |\det df|$ ) together imply for a typical  $x$

$$\lambda_1(\mu) + \dots + \lambda_{\dim M}(\mu) = \lim_{n \rightarrow \infty} \frac{1}{n} \log |\det df^n(x)| = \int \log |\det df| d\mu < +\infty. \quad (4)$$

Hence (3) and (4) imply  $\log \|(df)^{-1}\| \in L^1(\mu)$  and thus  $\log^+ \|(df)^{-1}\| \in L^1(\mu)$ . Since  $f$  is  $C^{1+\beta}$ , using for example (3) for any  $x$  sufficiently close to  $\mathcal{S}$  we obtain

$$m(df(x))^{\dim M} \leq |\det df(x)| \leq \text{Const} \cdot d(x, \mathcal{S})^\beta$$

Hence  $\log d(\cdot, \mathcal{S}) \in L^1(\mu)$ .  $\square$

**Example 2** (Continuous interval maps with flat or non-flat tops). Any  $C^{1+\beta}$  interval map, so in particular the quadratic family and any multimodal map are in the above setting. If  $f$  is a S-unimodal Misiurewicz map (that is, if  $f: [a, b] \rightarrow [a, b]$  is  $C^3$ , satisfies  $f(a) = f(b) = a$ , possesses a unique critical point  $c \in (a, b)$ , has non-positive Schwarzian derivative, and the critical point is non-recurrent and  $f$  has no sinks) then there exists a  $f$ -invariant absolutely continuous  $\sigma$ -finite Borel measure, that is finite if and only if  $\log |f'| \in L^1(m)$  [28] and in this case  $m$  has positive entropy and a positive Lyapunov exponent. Here the critical point can be either non-flat or flat, and we refer to [24] for an example of a  $C^\infty$  map with a  $C^\infty$  flat top for that  $\log |f'| \in L^1(m)$  is satisfied.

**Example 3** (Holomorphic maps). Consider a continuous map of the Riemann sphere that can be analytically extended to an open neighborhood of some compact set  $X \subset \overline{\mathbb{C}}$ . This includes the case of Julia sets of rational maps with (necessarily non-flat) critical points (see [21, 20]) for which any ergodic measure with positive entropy fits the hypotheses of Theorem 3.

**Example 4** (Skew-products of quadratic maps). Consider the following family of maps  $f: S^1 \times \mathbb{R} \rightarrow S^1 \times \mathbb{R}$

$$f(s, x) = (ds \mod 1, a - x^2 + \alpha \sin(2\pi s))$$

introduced by Viana in [26]. Here  $d \geq 2$  is an integer,  $\alpha \in \mathbb{R}$ , and  $a \in (1, 2)$  is such that the quadratic map  $g_a(x) = a - x^2$  has a pre-periodic (but not periodic) critical point. Viana [26] (for  $d \geq 16$ ) shows that Lebesgue almost every point possesses two positive Lyapunov exponents provided  $\alpha$  is sufficiently small. Alves [1] deduces that  $f$  possesses an absolutely continuous  $f$ -invariant Borel probability measure  $\mu$  that hence has only positive finite Lyapunov exponents. Buzzzi et. al [6] generalize these results to the case

$d \geq 2$ . In addition, note that by the Pesin formula [13, Theorem 1.1] the measure  $\mu$  has positive entropy. Hence, Theorem 3 applies to  $\mu$ .

## 2. PRELIMINARIES

We collect some preparatory results.

**2.1. Rokhlin natural extension.** The fundamental approach in obtaining the desired ergodic properties is to study a related invertible system that unravels the different preimages of a point. As our analysis is based on the asymptotic behavior of infinite orbits we need to exclude points that eventually are mapped onto  $\mathcal{S}$ . Set  $N^+ \stackrel{\text{def}}{=} \{x \in M : f^n(x) \notin \mathcal{S} \text{ for all } n \geq 0\}$ . Consider the set

$$N \stackrel{\text{def}}{=} \bigcap_{n \geq 0} f^n(N^+).$$

Note that  $N$  is invariant with respect to  $f$ , that is, satisfies  $f(N) = N$ . Given the transformation  $f : N \rightarrow N$  we consider the natural extension  $\hat{f} : \hat{N} \rightarrow \hat{N}$  given by  $\hat{f}(\dots, x_{-1}, x_0) = (\dots, x_{-1}, x_0, f(x_0))$  on

$$\hat{N} = \{\hat{x} = (x_{-n})_{n \geq 0} : f(x_{-n-1}) = x_{-n} \text{ for every } n \geq 0, x_n \in N\},$$

which is indeed an extension through the natural projection map  $\pi : \hat{N} \rightarrow N$  defined through  $\pi\hat{x} = x_0$ . The inverse map is given by  $\hat{f}^{-1}(\dots, x_{-1}, x_0) = (\dots, x_{-2}, x_{-1})$ . Giving  $\hat{N}$  the relative topology as a subset of the product  $N^{\mathbb{Z}_+}$ , we obtain a homeomorphism  $\hat{f}$  of  $\hat{N}$ . We equip  $\hat{N}$  with the metric  $d(\hat{x}, \hat{y}) = \sum_{k \geq 0} 2^{-k} d(x_{-k}, y_{-k})$ . Given an  $f$ -invariant ergodic Borel probability measure  $\mu$ , the unique measure  $\hat{\mu} \in \mathcal{M}(\hat{f})$  for which we have  $\pi_*\hat{\mu} = \mu$  satisfies  $h_\mu(f) = h_{\hat{\mu}}(\hat{f})$  [25]. If  $\mu$  is ergodic then so is  $\hat{\mu}$ , and  $\hat{\mu}$  is also ergodic and invariant with respect to  $\hat{f}^{-1}$ .

Given  $\hat{x} = (\dots, x_{-1}, x_0)$ , in the following we will use the notation  $f_{x_{-n}}^{-n}$  for the corresponding inverse branch of the map  $f|_{B(x_{-n}, \delta)} \circ \dots \circ f|_{B(x_0, \delta)}$  whenever  $\delta$  is chosen sufficiently small such that each of those maps  $f|_{B(\cdot, \delta)}$  is invertible.

**2.2. Slow recurrence towards the set  $\mathcal{S}$ .** Although no trajectory in  $N$  ever hits the set  $\mathcal{S}$ , it may approach  $\mathcal{S}$  arbitrarily closely and hence the behavior of nearby trajectories may be difficult to control. However, this is ruled out under the assumption  $(C_2)$  as we show now.

We first provide some preliminary result. Given  $\delta > 0$ , denote  $B(\mathcal{S}, \delta) \stackrel{\text{def}}{=} \bigcup_{x \in \mathcal{S}} B(x, \delta)$ .

**Lemma 1.** *If  $\log d(\cdot, \mathcal{S}) \in L^1(\mu)$ , then  $\sum_{n \geq 1} \mu(B(\mathcal{S}, e^{-n\delta})) < +\infty$ .*

*Proof.* Given  $\delta > 0$ , observe that

$$\begin{aligned} 0 \leq \sum_{n \geq 1} \mu(B(\mathcal{S}, e^{-n\delta})) &= \sum_{n \geq 1} n \mu(B(\mathcal{S}, e^{-n\delta}) \setminus B(\mathcal{S}, e^{-(n+1)\delta})) \\ &\leq -\frac{1}{\delta} \sum_{n \geq 1} -n\delta \mu(B(\mathcal{S}, e^{-n\delta}) \setminus B(\mathcal{S}, e^{-(n+1)\delta})) \\ &\leq -\frac{1}{\delta} \int_{B(\mathcal{S}, e^{-\delta})} \log d(x, \mathcal{S}) d\mu(x). \end{aligned}$$

This proves the lemma.  $\square$

Lemma 1 implies  $\mu(\mathcal{S}) = 0$  (in fact, this follows already from  $\log d(\cdot, \mathcal{S}) \in L^1(\mu)$ ), and from  $f$ -invariance of the measure we conclude that  $\mu(N) = 1$ .

A typical backward branch of a point does not come too close to the set  $\mathcal{S}$  in the following sense.

**Lemma 2.** *If  $\log d(\cdot, \mathcal{S}) \in L^1(\mu)$ , then for any  $\delta > 0$  there exists a  $\hat{\mu}$ -full measure set  $\hat{\Lambda} \subset \hat{N}$  such that for every  $\hat{x} \in \hat{\Lambda}$  for only finitely many  $k \geq 1$  we have  $\pi \hat{f}^{-k}(\hat{x}) \in B(\mathcal{S}, e^{-k\delta})$ .*

*Proof.* By working in the inverse limit space, from  $\hat{f}$ -invariance of  $\hat{\mu}$  and from  $\pi_* \hat{\mu} = \mu$  we can conclude that

$$\sum_{k \geq 1} \hat{\mu}(\hat{f}^k \circ \pi^{-1}(B(\mathcal{S}, e^{-k\delta}))) = \sum_{k \geq 1} \mu(B(\mathcal{S}, e^{-k\delta})) < \infty$$

using Lemma 1. The claim now follows from the Borel-Cantelli lemma.  $\square$

**2.3. Lyapunov exponents.** We now consider Lyapunov exponents. While in positive time direction there is no change to define the largest (positive) exponent  $\bar{\lambda}$ , we need to change the definition of the smallest exponent  $\underline{\lambda}$  to handle non-invertibility of  $f$ : for given  $x \in M \setminus \mathcal{S}$  let

$$\bar{\lambda}(x) \stackrel{\text{def}}{=} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \|df^n(x)\|, \quad \underline{\lambda}(x) \stackrel{\text{def}}{=} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \|df^n(x)^{-1}\|^{-1}.$$

(If  $f$  is invertible this definition coincides with the usual one.)

If the measure  $\mu$  is ergodic then the Lyapunov exponents of the derivative cocycle with generator  $df(\pi \hat{x})$  coincide for  $\hat{\mu}$ -almost every  $\hat{x}$  with the Lyapunov exponents of  $\mu$ . In particular, if we consider the set of points that are Lyapunov regular with respect to  $f$ , then for every  $v \in T_x M$  the Lyapunov exponent of  $(x, v)$  we have

$$\lambda(x, v) \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} \log \|df^n(x)(v)\|,$$

satisfies  $\underline{\lambda}(x) \leq \lambda(x, v) \leq \bar{\lambda}(x)$ . Moreover, for  $\mu$ -almost every  $x$  and every  $v \in T_x M$

$$\lambda(x, v) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|(df(\pi \hat{f}^{-n}(\hat{x})))^{-1} \cdots (df(\pi \hat{f}^{-1}(\hat{x})))^{-1} (df(\pi \hat{x}))^{-1}(v)\|$$

and

$$\chi(\mu) \stackrel{\text{def}}{=} \int \log \|(df)^{-1}\|^{-1} d\mu \leq \lambda(x, v) \leq \int \log \|df\| d\mu.$$

**2.4. Lyapunov change of coordinates.** We now apply nowadays standard methods from Oseledets-Pesin theory by introducing a Lyapunov change of coordinates and appropriately chosen tempered sequences. We will largely follow arguments e.g. in [4, 16, 18, 21] without always giving any other particular reference.

Recall that a measurable function  $h: \widehat{N} \rightarrow \mathbb{R}$  is said to be *tempered* on  $\widehat{\Lambda} \subset \widehat{N}$  with respect to  $\widehat{f}$  if for every  $\widehat{x} \in \widehat{\Lambda}$

$$\lim_{k \rightarrow \pm\infty} \frac{1}{k} \log h(\widehat{f}^k(\widehat{x})) = 0.$$

We will use the following preliminary result on tempered sequences. For completeness we provide its proof (see for example [18, p. 293] and [10] for a related result).

**Lemma 3** (Tempering kernel lemma). *Given a positive measurable function  $\tilde{r}: \widehat{N} \rightarrow \mathbb{R}$  that is tempered on  $\widehat{\Lambda} \subset \widehat{N}$  with respect to  $\widehat{f}$  and  $\varepsilon > 0$ , there exists  $\widehat{\Gamma} \subset \widehat{\Lambda}$  with  $\widehat{\mu}(\widehat{\Gamma}) = 1$  and a positive measurable function  $r$  on  $\widehat{\Gamma}$  satisfying  $0 < r \leq \tilde{r}$  and*

$$\frac{r(\widehat{x})}{r(\widehat{f}^k(\widehat{x}))} \leq e^{|k|\varepsilon} \quad \text{for every } k \in \mathbb{Z} \text{ and every } \widehat{x} \in \widehat{\Gamma}.$$

*Proof.* Because  $\tilde{r}$  is tempered, given  $\varepsilon > 0$  for each  $\widehat{x} \in \widehat{\Lambda}$  there are constants  $c_1, c_2 > 0$  so that  $c_1 e^{-|k|\varepsilon} \leq \tilde{r}(\widehat{f}^k(\widehat{x})) \leq c_2 e^{|k|\varepsilon}$  for every  $k \in \mathbb{Z}$ . Let

$$b(\widehat{x}) \stackrel{\text{def}}{=} \inf_{n \geq 0} \tilde{r}(\widehat{f}^n(\widehat{x})) e^{n\varepsilon}.$$

Observe that  $0 < b(\widehat{x}) \leq \tilde{r}(\widehat{x})$  and that

$$b(\widehat{f}(\widehat{x})) = \inf_{n \geq 1} \tilde{r}(\widehat{f}^{n+1}(\widehat{x})) e^{n\varepsilon} = e^{-\varepsilon} \inf_{n \geq 1} \tilde{r}(\widehat{f}^{n+1}(\widehat{x})) e^{(n+1)\varepsilon} \geq e^{-\varepsilon} b(\widehat{x})$$

for every  $\widehat{x} \in \widehat{\Lambda}$ . Hence

$$\log b(\widehat{f}(\widehat{x})) - \log b(\widehat{x}) \geq -\varepsilon, \tag{5}$$

and [16, Lemma III.8] implies that  $b$  is tempered with respect to  $\widehat{f}$  on a full measure subset  $\widehat{\Gamma} \subset \widehat{\Lambda}$ . Hence, for every  $\widehat{x} \in \widehat{\Gamma}$  there exists  $c > 0$  so that  $b(\widehat{f}^{-n}(\widehat{x})) \geq c e^{-n\varepsilon}$  for every  $n \geq 0$ . Now for every  $\widehat{x} \in \widehat{\Gamma}$  let

$$r(\widehat{x}) \stackrel{\text{def}}{=} \inf_{n \geq 0} b(\widehat{f}^{-n}(\widehat{x})) e^{n\varepsilon}.$$

Observe that

$$r(\widehat{f}^{-1}(\widehat{x})) = e^{-\varepsilon} \inf_{n \geq 1} b(\widehat{f}^{-n}(\widehat{x})) e^{n\varepsilon} \geq e^{-\varepsilon} r(\widehat{x})$$



for every  $\hat{x} \in \hat{\Gamma}$ . Since  $\hat{f}$  is invertible, hence we have

$$\log r(\hat{f}(\hat{x})) - \log r(\hat{x}) \geq -\varepsilon$$

for every  $\hat{x} \in \hat{\Gamma}$ . Thus [16, Lemma III.8] implies that  $r$  is tempered with respect to  $\hat{f}$  on a full measure subset of  $\hat{\Gamma}$ . This proves the lemma.  $\square$

We denote by  $\|\cdot\|$  the norm on  $T_x M$  induced by the Riemannian metric on  $M$ . The set  $N \subset M \setminus \mathcal{S}$  is invariant with respect to  $f$  and so we can consider the restriction of the tangent map  $df$  to  $T_N M$ . Let us denote  $T(\hat{x}) \stackrel{\text{def}}{=} df(\pi\hat{x})$ ,  $T_{\hat{x}}^n \stackrel{\text{def}}{=} T(\hat{f}^{n-1}(\hat{x})) \cdots T(\hat{f}(\hat{x}))T(\hat{x})$ , and

$$T_{\hat{x}}^{-n} \stackrel{\text{def}}{=} T^{-1}(\hat{f}^{-n}(\hat{x})) \cdots T^{-1}(\hat{f}^{-2}(\hat{x}))T^{-1}(\hat{f}^{-1}(\hat{x})).$$

Under our hypothesis  $(C_1)$  we have  $\log^+ \|T\|, \log^+ \|T^{-1}\| \in L^1(\hat{\mu})$ . Thus the assumptions of the multiplicative ergodic theorem applied to  $(\hat{N}, \hat{f}, \hat{\mu})$  are met (see, for example, [22, Theorem 3.1]).

**Lemma 4.** *Given  $\varepsilon \in (0, \chi/3)$ , there exists a compact set  $\hat{\Lambda} \subset \hat{N}$  of full measure and a function  $C_\varepsilon$  that is tempered on  $\hat{\Lambda}$  with respect to  $f$  such that for every  $\hat{x} \in \hat{\Lambda}$ ,  $v \in T_{\pi\hat{x}} M$ , and  $n \geq 1$  we have*

$$\|T_{\hat{x}}^{-n}(v)\| \leq C_\varepsilon(\hat{x}) e^{-n(\chi-\varepsilon)} \|v\|. \quad (6)$$

*Proof.* It follows that there exists a set  $\hat{\Lambda} \subset \hat{N}$  of full measure, a positive integer  $s \leq \dim M$  and numbers  $\chi_1 < \cdots < \chi_s$  such that for every  $\hat{x} \in \hat{\Lambda}$  and  $v \in T_x M \setminus \{0\}$  the limit

$$\chi(\hat{x}, v) \stackrel{\text{def}}{=} \lim_{k \rightarrow \pm\infty} \frac{1}{k} \log \|T_{\hat{x}}^k(v)\|$$

exists and is equal to one of the numbers  $\chi_i$ ,  $i = 1, \dots, s$ . Moreover, for  $\hat{\mu}$ -almost every  $\hat{x}$  we have  $\chi(\hat{x}, v) = \lambda(\pi\hat{x}, v)$ . By our assumption,

$$0 < \chi \stackrel{\text{def}}{=} \chi(\mu) \leq \chi(\hat{x}, v).$$

Given any  $0 < \varepsilon < \chi/3$ , for every  $\hat{x} \in \hat{\Lambda}$  and  $v \in T_{\pi\hat{x}} M$  we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|T_{\hat{x}}^{-n}(v)\| < -\chi + \varepsilon.$$

Thus there is a measurable function  $C_\varepsilon : \hat{\Lambda} \rightarrow \mathbb{R}$  given by

$$C_\varepsilon(\hat{x}) \stackrel{\text{def}}{=} \sup_{n \geq 0} \|T_{\hat{x}}^{-n}\| e^{n(\chi-\varepsilon)} < \infty.$$

In particular,  $C_\varepsilon$  is also well-defined at every  $\hat{f}^{-k}(\hat{x})$ ,  $k \geq 1$ . Observe that

$$\begin{aligned} C_\varepsilon(\hat{f}(\hat{x})) &\leq \sup \left\{ 1, \sup_{n \geq 1} \|T_{\hat{f}(\hat{x})}^{-n}\| e^{n(\chi-\varepsilon)} \right\} \\ &\leq \sup \left\{ 1, \sup_{n \geq 1} \|T_{\hat{x}}^{-n-1}\| \|T_{\hat{f}(\hat{x})}^{-1}\| e^{n(\chi-\varepsilon)} \right\} \\ &\leq \sup \left\{ 1, e^{-\chi+\varepsilon} \|T_{\hat{f}(\hat{x})}^{-1}\| \sup_{n \geq 1} \|T_{\hat{x}}^{-n-1}\| e^{(n+1)(\chi-\varepsilon)} \right\} \\ &= \sup \left\{ 1, e^{-\chi+\varepsilon} \|T_{\hat{f}(\hat{x})}^{-1}\| C_\varepsilon(\hat{x}) \right\} \leq C_\varepsilon(\hat{x}) \sup \left\{ 1, e^{-\chi+\varepsilon} \|T_{\hat{f}(\hat{x})}^{-1}\| \right\}. \end{aligned}$$

We obtain

$$\log C_\varepsilon(\hat{f}(\hat{x})) - \log C_\varepsilon(\hat{x}) \leq \log \sup \left\{ 1, e^{-\chi+\varepsilon} \|T_{\hat{f}(\hat{x})}^{-1}\| \right\},$$

which is bounded from above. This justifies that  $\log C_\varepsilon \circ \hat{f} - \log C_\varepsilon \in L^1(\hat{\mu})$  and hence that  $C_\varepsilon$  is tempered on  $\hat{\Lambda}$ , according to [16, Lemma III.8].

Given  $\varepsilon > 0$ ,  $\hat{x} \in \hat{\Lambda}$ , and  $v \in T_{\pi\hat{x}}M$ , let us introduce a Lyapunov change of coordinates  $\|v\| \mapsto \|v\|'_\hat{x}$  by

$$\|v\|'_\hat{x} \stackrel{\text{def}}{=} \sup_{n \geq 0} \|T_{\hat{x}}^{-n}(v)\| e^{n(\chi-\varepsilon)}.$$

Notice that for every  $\hat{x}$  we have  $C_\varepsilon(\hat{x}) \geq 1$  and

$$\|v\| \leq \|v\|'_\hat{x} \leq C_\varepsilon(\hat{x}) \|v\|. \quad (7)$$

Further, we have

$$\begin{aligned} \|T_{\hat{x}}^{-1}(v)\|'_{\hat{f}^{-1}(\hat{x})} &= \sup_{n \geq 0} \|T_{\hat{f}^{-1}(\hat{x})}^{-n} T_{\hat{x}}^{-1}(v)\| e^{n(\chi-\varepsilon)} \\ &= e^{-\chi+\varepsilon} \left( \sup_{n \geq 0} \|T_{\hat{x}}^{-n-1}(v)\| e^{(n+1)(\chi-\varepsilon)} \right) \\ &\leq e^{-\chi+\varepsilon} \|v\|'_\hat{x}, \end{aligned} \quad (8)$$

which implies that for every  $n \geq 1$

$$\|T_{\hat{x}}^{-n}(v)\|'_{\hat{f}^{-n}(\hat{x})} \leq e^{-n(\chi-\varepsilon)} \|v\|'_\hat{x}$$

Using (7), this implies (6). The lemma is proved.  $\square$

**2.5. Construction of local unstable manifolds.** Non-invertibility of  $f$  implies that points not necessarily have unique local unstable manifolds. However, given  $R > 0$  and  $x \in N$ , based on the natural extensions we can study the following type of sets

$$\{y \in N : \exists(\dots, y_{-1}, y) = \hat{y} \in \hat{N} \forall k \geq 0 : d(x_{-k}, y_{-k}) < R\}.$$

Related constructions of unstable manifolds have been introduced by Ledrappier [12] in the case of piecewise  $C^{1+\beta}$  interval maps. They can also be read from Newhouse [18] for  $C^{1+\beta}$  endomorphisms in the higher-dimensional case.

The case of  $C^2$  maps with singular points under slightly stronger conditions was also covered in [11] and the holomorphic case in [21].

We will use the following result.

**Lemma 5.** *Given  $\varepsilon \in (0, \chi/2)$  and  $\delta \in (0, 1)$ , there exist a compact set  $\widehat{\Lambda}_1 = \widehat{\Lambda}_1(\varepsilon, \delta) \subset \widehat{\Lambda}$  and a number  $\rho = \rho(\varepsilon, \delta) > 0$  such that  $\widehat{\mu}(\widehat{\Lambda}_1) > 1 - \delta$  and that for every  $\widehat{x} \in \widehat{\Lambda}_1$ ,  $x = \pi\widehat{x}$ ,  $y \in B(x, \rho)$ , and  $k \geq 1$  we have*

$$\|df_{x-k}^{-k}(y)\| \leq e^{-k(\chi-2\varepsilon)}. \quad (9)$$

In particular

$$d(f_{x-k}^{-k}(x), f_{x-k}^{-k}(y)) \leq e^{-k(\chi-2\varepsilon)} d(x, y).$$

*Proof.* Given points  $x, y \in M \setminus \mathcal{S}$ , a vector  $v \in T_x M$ , and points  $\widehat{x}, \widehat{y} \in \widehat{N}$  with  $\pi\widehat{x} = x$ ,  $\pi\widehat{y} = y$ , then by our Hölder assumption ( $C_2$ ) we have

$$\|T_{\widehat{x}}^{-1}(v) - T_{\widehat{y}}^{-1}(v)\| \leq H(x) d(x, y)^\beta \|v\|$$

whenever  $d(x, y) < G(x)$ . For notational simplicity, let us refrain from considering the length change of a given vector  $v \in T_x M$  by changing between charts, which can be made arbitrarily small by shrinking the domain of charts. When  $d(\widehat{x}, \widehat{y})$  is small, let us define  $\|v\|_{\widehat{y}}' = \|v\|_{\widehat{x}}'$  for  $v \in T_{\pi\widehat{x}} M$ . We then obtain from the triangle inequality and from (8) and (7)

$$\begin{aligned} \|T_{\widehat{y}}^{-1}(v)\|_{\widehat{f}^{-1}(\widehat{y})}' &= \|T_{\widehat{y}}^{-1}(v)\|_{\widehat{f}^{-1}(\widehat{x})}' \\ &\leq \|T_{\widehat{x}}^{-1}(v)\|_{\widehat{f}^{-1}(\widehat{x})}' + \|T_{\widehat{y}}^{-1}(v) - T_{\widehat{x}}^{-1}(v)\|_{\widehat{f}^{-1}(\widehat{x})}' \\ &\leq e^{-\chi+\varepsilon} \|v\|_{\widehat{x}}' + C_\varepsilon(\widehat{f}^{-1}(\widehat{x})) \|T_{\widehat{y}}^{-1}(v) - T_{\widehat{x}}^{-1}(v)\| \\ &\leq e^{-\chi+\varepsilon} \|v\|_{\widehat{x}}' + C_\varepsilon(\widehat{f}^{-1}(\widehat{x})) H(\pi\widehat{x}) d(\pi\widehat{x}, \pi\widehat{y})^\beta \|v\|_{\widehat{x}}' \end{aligned} \quad (10)$$

Let

$$\widetilde{r}(\widehat{x}) \stackrel{\text{def}}{=} \min \left\{ \left( \frac{C_\varepsilon(\widehat{x})^{-1} e^{-\chi+2\varepsilon} - e^{-\chi+\varepsilon}}{C_\varepsilon(\widehat{f}^{-1}(\widehat{x})) H(\pi\widehat{x})} \right)^{1/\beta}, 1, G(\pi\widehat{x}) \right\}.$$

Recall that  $C_\varepsilon$  is tempered on  $\widehat{\Lambda}$ . From  $\log H \in L^1(\mu)$  we conclude  $\log H \circ \pi \in L^1(\widehat{\mu})$  and hence that  $H \circ \pi$  is tempered on a full measure subset of  $\widehat{\Lambda}$ . Analogously, we conclude that  $G \circ \pi$  and hence that  $\widetilde{r}$  is tempered on a full measure subset of  $\widehat{\Lambda}$ . Lemma 3 implies that there exists a full-measure subset  $\widehat{\Lambda}_1 \subset \widehat{\Lambda}$  and a positive measurable function  $\widehat{x} \mapsto r(\widehat{x})$  defined on  $\widehat{\Lambda}_1$  such that for every  $\widehat{x} \in \widehat{\Lambda}_1$  we have  $0 < r(\widehat{x}) \leq \widetilde{r}(\widehat{x})$  and for every  $k \in \mathbb{Z}$

$$\frac{r(\widehat{x})}{r(\widehat{f}^k(\widehat{x}))} \leq e^{|k|\varepsilon}. \quad (11)$$

We hence obtain for every  $\widehat{x} \in \widehat{\Lambda}_1$  and every  $\widehat{y}$  satisfying  $d(\pi\widehat{x}, \pi\widehat{y}) < r(\widehat{x})$

$$\|T_{\widehat{y}}^{-1}(v)\| \leq e^{-\chi+2\varepsilon} \|v\|$$

and in particular together with (11) also

$$d(\pi \hat{f}^{-1}(\hat{x}), \pi \hat{f}^{-1}(\hat{y})) \leq e^{-\chi+2\varepsilon} d(\pi \hat{x}, \pi \hat{y}) < e^{-\chi+3\varepsilon} e^{-\varepsilon} r(\hat{x}) \leq r(\hat{f}^{-1}(\hat{x})).$$

After this preparation of distortion control for one single iteration, we want to achieve uniform contraction on an entire backward branch. Removing at most a set of points of zero measure, we may assume that for every  $\hat{x} \in \hat{\Lambda}_1$  the above statements and the statements of Lemma 4 and Lemma 2 applied to  $\delta = \chi$  are true. That is, in particular, for every point  $\hat{x} = (\dots, x_{-1}, x_0) \in \hat{\Lambda}$  there exists a number  $m(\hat{x}) \geq 1$  such that for every  $k \geq m(\hat{x})$

$$\|T_{\hat{x}}^{-k}(v)\| \leq e^{-k(\chi-2\varepsilon)} \|v\|$$

and that

$$x_{-k} \notin B(\mathcal{S}, e^{-k\chi}).$$

We can now choose a positive measurable function  $R: \hat{\Lambda} \rightarrow \mathbb{R}$  such that for every  $0 \leq k \leq m(\hat{x})$  all the inverse branches  $f_{x_{-k}}^{-k}: B(x_0, r(\hat{x})R(\hat{x})) \rightarrow M$  are well-defined, that

$$\text{diam}\left(f_{x_{-k}}^{-k}(B(x_0, r(\hat{x})R(\hat{x})))\right) < \min\left\{e^{-k\chi}, r(\hat{f}^{-k}(\hat{x}))\right\},$$

and that distortion is bounded in the sense that for every  $0 \leq k \leq m(\hat{x})$

$$\max \frac{\|T_{\hat{y}_1}^{-k}(v)\|}{\|T_{\hat{y}_2}^{-k}(v)\|} \leq e^{k\varepsilon} C_\varepsilon(\hat{x})^{-1}, \quad (12)$$

where the maximum is taken over all  $\hat{y}_i$  with  $d(\pi \hat{y}_i, \pi \hat{x}) \leq r(\hat{x})R(\hat{x})$  and all  $v \in T_{\hat{x}}M$ . Thus, applying (12) and then (6), we obtain for every  $y \in B(x_0, r(\hat{x})R(\hat{x}))$  and also every  $0 \leq k \leq m(\hat{x})$

$$\|df_{x_{-k}}^{-k}(y)\| \leq e^{k\varepsilon} C_\varepsilon(\hat{x})^{-1} \|df_{x_{-k}}^{-k}(x)\| \leq e^{k\varepsilon} e^{-k(\chi-\varepsilon)} = e^{-k(\chi-2\varepsilon)}.$$

In particular, for  $m = m(\hat{x})$

$$d(f_{x_{-m}}^{-m}(y), x_{-m}) \leq r(x_{-m}) \leq G(x_{-m})$$

and  $f_{x_{-m}}^{-m}: B(x_0, e^{-m\chi}) \rightarrow M$  is well-defined. After making such choices, we are in the above setting of distortion control and given  $y \in B(x_0, r(\hat{x})R(\hat{x}))$ , we have

$$d(f_{x_{-(m+1)}}^{-(m+1)}(y), x_{-(m+1)}) \leq r(x_{-(m+1)}) \leq G(x_{-(m+1)}).$$

By induction, we can now conclude that (9) is true for every  $\hat{x} \in \hat{\Lambda}_1$ , for every  $y \in B(\pi \hat{x}, r(\hat{x})R(\hat{x}))$  and every  $k \geq 1$ .

The claimed properties now follow from the Lusin theorem.  $\square$

**2.6. Uniform recurrence.** Let  $\mathcal{P}$  be any finite measurable partition of  $M$ . Denote by  $\mathcal{P}(x)$  the partition element that contains  $x$ . Let  $\varphi_1, \dots, \varphi_K: M \rightarrow \mathbb{R}$  be continuous functions. The following fact is an immediate consequence of the Birkhoff ergodic theorem.

**Lemma 6.** *Given numbers  $\varepsilon > 0$  and  $\delta > 0$  and a positive measure set  $\widehat{\Lambda}_1 \subset \widehat{N}$ , there exist a positive integer  $n_2 = n_2(\varepsilon, \delta)$  and a compact set  $\widehat{\Lambda}_2 = \widehat{\Lambda}_2(\varepsilon, \delta)$  such that  $\widehat{\mu}(\widehat{\Lambda}_2) > 1 - \delta$  so that for every  $\widehat{x} \in \widehat{\Lambda}_2$  and every  $n \geq n_2$  we have*

a)  $\pi \widehat{f}^k(\widehat{x}) \in \mathcal{P}(\pi \widehat{x})$  and  $\widehat{f}^k(\widehat{x}) \in \widehat{\Lambda}_1$  for some number  $k \in [n, n + \varepsilon n]$ ,

b) for  $x = \pi \widehat{x}$  for every  $i = 0, \dots, K$  and every  $k \geq n$  we have

$$\left| \frac{1}{k} \left( \varphi_i(x) + \varphi_i(f^1(x)) + \dots + \varphi_i(f^{k-1}(x)) \right) - \int \varphi_i d\mu \right| \leq \varepsilon.$$

*Proof.* Suppose that the partition  $\mathcal{P}$  has  $j$  elements  $\mathcal{P} = \{P_1, \dots, P_j\}$ . This partition of  $M$  naturally induces a partition  $\widehat{\mathcal{P}} = \{\widehat{P}_1, \dots, \widehat{P}_j\}$  of  $\widehat{N}$  given by  $\widehat{P}_i = \{\widehat{x}: \pi \widehat{x} \in P_i\}$ . Let  $\kappa \stackrel{\text{def}}{=} \min\{\varepsilon, \mu(\widehat{\Lambda}_1 \cap \widehat{P}_i)/4: i = 1, \dots, j\}$ . From the Birkhoff ergodic theorem we derive that for each  $i = 1, \dots, j$  for  $\widehat{\mu}$ -almost every  $\widehat{x}$  we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{card}\{k \in \{0, \dots, n-1\}: \widehat{f}^k(\widehat{x}) \in \widehat{\Lambda}_1 \cap \widehat{P}_i\} = \widehat{\mu}(\widehat{\Lambda}_1 \cap \widehat{P}_i)$$

Using the Eg'orov theorem and the Lusin theorem, we can conclude that there exists a compact set  $\widehat{\Lambda}_2 \subset \widehat{N}$  of measure  $\geq 1 - \delta$  such that the convergence is uniform on  $\widehat{\Lambda}_2$ . Hence, we find a number  $n_2 \geq 1$  such that for every  $i = 0, \dots, j$ , every  $\widehat{x} \in \widehat{\Lambda}_2$ , and every  $n \geq n_2$  we have

$$\left| \text{card}\{k \in \{0, \dots, n-1\}: \widehat{f}^k(\widehat{x}) \in \widehat{\Lambda}_1 \cap \widehat{P}_i\} - \mu(\widehat{\Lambda}_1 \cap \widehat{P}_i)n \right| \leq \kappa^2 n.$$

Assume that  $n_2$  is chosen large enough that  $\min_{i=1, \dots, j} (\mu(\widehat{\Lambda}_1 \cap \widehat{P}_i) - 3\kappa)n_2\varepsilon > 1$ . Thus, for every  $\widehat{x} \in \widehat{\Lambda}_2$ , every  $i = 0, \dots, j$ , and every  $n \geq n_2$  we obtain

$$\begin{aligned} & \text{card}\{k \in \{n, \dots, n(1 + \varepsilon) - 1\}: \widehat{f}^k(\widehat{x}) \in \widehat{\Lambda}_1 \cap \widehat{P}_i\} \\ & \geq \mu(\widehat{\Lambda}_1 \cap \widehat{P}_i)n(1 + \varepsilon) - n(1 + \varepsilon)\kappa^2 - \mu(\widehat{\Lambda}_1 \cap \widehat{P}_i)(n-1) - (n-1)\kappa^2 \\ & \geq n\varepsilon(\mu(\widehat{\Lambda}_1 \cap \widehat{P}_i) - \kappa^2) - 2n\kappa^2 \\ & \geq n\varepsilon(\mu(\widehat{\Lambda}_1 \cap \widehat{P}_i) - 3\kappa) > 1. \end{aligned}$$

Clearly,  $f^k(\pi \widehat{x}) \in P_i$  if  $\widehat{f}^k(\widehat{x}) \in \widehat{P}_i$ . In particular, this is true for the index  $i$  with  $P_i = \mathcal{P}(\pi \widehat{x})$ . This proves a).

Similar arguments apply to Birkhoff averages of the continuous function  $\varphi_1, \dots, \varphi_K$  that guarantee that the above applies and at the same time Birkhoff averages are uniformly converging on  $\widehat{\Lambda}_2$ . This proves the lemma.  $\square$

## 3. PROOF OF THEOREM 1

To construct a repeller on which the topological entropy  $f$  is roughly equal to  $h_\mu(f)$  (and on which the other required dynamic properties are also satisfied), we will produce a sufficiently large number of points that have distinguishable orbits of a certain length. First, recall that by [9, Theorem 1.1] we have

$$h_\mu(f) = \lim_{\tilde{\varepsilon} \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log N(\mu, n, \tilde{\varepsilon}),$$

where  $N(\mu, n, \tilde{\varepsilon})$  denotes the minimal number of sets

$$B_n(x, \tilde{\varepsilon}) \stackrel{\text{def}}{=} \left\{ y : d(f^k(x), f^k(y)) \leq \tilde{\varepsilon} \text{ for every } 0 \leq k \leq n-1 \right\}$$

that are needed to cover a set of measure greater than  $1 - \delta$ . Here  $\delta \in (0, 1)$  is an arbitrary number. Hence, fixing some  $\delta \in (0, 1)$  and some number  $\varepsilon > 0$  there exists  $\tilde{\varepsilon}_0 = \tilde{\varepsilon}_0(\varepsilon, \delta) > 0$  such that for every  $\tilde{\varepsilon} \leq \tilde{\varepsilon}_0$  there exists a number  $n_0 = n_0(\tilde{\varepsilon}, \varepsilon, \delta)$  satisfying the following. Given a set  $A$  of measure  $> 1 - 2\delta$  and a number  $n \geq n_0$ , then any  $(n, \tilde{\varepsilon})$ -separated set  $E \subset A$  of maximal cardinality satisfies

$$\log \text{card } E \geq n(h_\mu(f) - \varepsilon). \quad (13)$$

Let us start by choosing some  $\varepsilon \in (0, \chi/3)$  and let  $\tilde{\varepsilon}_0 = \tilde{\varepsilon}_0(\varepsilon, \delta)$ .

Those points in  $E$  with distinguishable orbits in the following will be placed in some Lyapunov regular set and hence additionally will have uniformly hyperbolic behavior. By the preparatory results in Lemmas 4 and 5 there exist a number  $\rho = \rho(\varepsilon/3, \delta) > 0$  and a compact set  $\hat{\Lambda}_1 = \hat{\Lambda}_1(\varepsilon/3, \delta) \subset \hat{\Lambda}$  of measure  $> 1 - \delta$  such that for every  $\hat{x} \in \hat{\Lambda}_1$  and every  $y \in B(\pi\hat{x}, \rho)$ , for all  $k \geq 1$  we have

$$\|df_{x_{-k}}^{-k}(y)\| \leq e^{-k(\chi-2\varepsilon)} \quad (14)$$

and in particular

$$d(f_{x_{-k}}^{-k}(y), x_{-k}) \leq e^{-k(\chi-2\varepsilon)} d(\pi\hat{x}, y). \quad (15)$$

Let  $\Lambda_1 \stackrel{\text{def}}{=} \pi\hat{\Lambda}_1$  and note that  $\mu(\Lambda_1) > 1 - \delta$ .

We cover  $\Lambda_1$  by balls

$$B(x_1, \rho/2), \dots, B(x_j, \rho/2) \quad (16)$$

that are centered at points  $x_i \in \Lambda_1$ ,  $i = 1, \dots, j$ . We choose a number  $n_1 = n_1(j, \varepsilon)$  satisfying

$$n_1 \geq \frac{\log j}{\varepsilon}. \quad (17)$$

Without loss of generality we can assume that  $\rho > 0$  was chosen so small that for every  $i = 1, \dots, K$ , for every  $x$  and every  $y \in B(x, \rho)$  we have

$$|\varphi_i(x) - \varphi_i(y)| \leq \varepsilon. \quad (18)$$

Besides the Lyapunov regular cover (16), let us consider a finite partition  $\mathcal{P}$  of  $\Lambda$  of diameter  $< \rho/4$ . Notice that each partition element  $\mathcal{P}(x_i)$ ,  $i = 1, \dots, j$ , satisfies  $\mathcal{P}(x_i) \subset B(x_i, \rho/2)$ .

We also want those points in  $E$  with distinguishable orbits, or at least most of them, in addition to be closely recurring to itself at the same, or at least almost the same, time. By Lemma 6 a), there exists a number  $n_2 = n_2(\varepsilon, \delta)$  and a compact set  $\widehat{\Lambda}_2 = \widehat{\Lambda}_2(\varepsilon, \delta) \subset \widehat{N}$  of points of measure  $> 1 - \delta$  such that for every  $n \geq n_2$  and for every  $\widehat{x} \in \widehat{\Lambda}_2$  we have

$$f^k(\pi\widehat{x}) = \pi\widehat{f}^k(\widehat{x}) \in \mathcal{P}(\pi\widehat{x}) \quad \text{and} \quad \widehat{f}^k(\widehat{x}) \in \widehat{\Lambda}_1 \quad (19)$$

for some  $k \in [n, n + n\varepsilon]$ . We choose a number  $n_3 = n_3(\varepsilon, \chi)$  satisfying

$$n_3 \geq \frac{\log 4}{\chi - 2\varepsilon}. \quad (20)$$

Choose  $\widetilde{\varepsilon} \in (0, \min\{\widetilde{\varepsilon}_0, \rho/2\})$  and notice that any  $(n, \rho/2)$ -separated set is also  $(n, \widetilde{\varepsilon})$ -separated. Let  $n_0 = n_0(\widetilde{\varepsilon}, \varepsilon, \delta)$ .

We take

$$n \geq \max\{n_0, n_1, n_2, n_3\} \quad (21)$$

and consider the set

$$\widehat{\Lambda}_* \stackrel{\text{def}}{=} \widehat{\Lambda}_1 \cap \widehat{\Lambda}_2.$$

Let  $\Lambda_* = \pi\widehat{\Lambda}_*$ . Notice that  $\widehat{\mu}(\widehat{\Lambda}_*) \geq 1 - 2\delta$  and hence  $\mu(\Lambda_*) > 1 - 2\delta$ .

We now choose an  $(n, \rho/2)$ -separated set  $E \subset \Lambda_*$  that is of maximal cardinality. Note that then  $E$  is also  $(n, \rho/2)$ -spanning and hence  $\bigcup_{x \in E} B_n(x, \rho/2)$  covers the set  $\Lambda_*$  that has measure  $> 1 - 2\delta$ .

Remembering (19), we now partition the set  $E$  into sets  $F_k$ ,  $n \leq k < n + \varepsilon n$ , defined by

$$F_k \stackrel{\text{def}}{=} \left\{ x \in E : f^k(x) \in \mathcal{P}(x) \right\},$$

that is, having the same time  $k$  of return to their partition element. Let  $m$  be the index satisfying  $\text{card } F_m = \max_{n \leq k < n + \varepsilon n} \text{card } F_k$ . Since  $\text{card } E = \sum_{n \leq k < n + \varepsilon n} \text{card } F_k$ , we have  $\varepsilon n \text{card } F_m \geq \text{card } E$ . With  $\varepsilon n < e^{n\varepsilon}$  and (13) we obtain

$$\text{card } F_m \geq \frac{\text{card } E}{\varepsilon n} \geq e^{n(h_\mu(f) - 2\varepsilon)}.$$

In the following we will consider only the ball  $B(x_i, \rho/2)$  from the cover (16) for that  $\text{card}(F_m \cap \mathcal{P}(x_i))$  is maximal. Hence we have

$$\text{card}(F_m \cap \mathcal{P}(x_i)) \geq \frac{1}{j} \text{card } F_m \geq \frac{1}{j} e^{n(h_\mu(f) - 2\varepsilon)}. \quad (22)$$

Recall that exactly after  $m$  iterations each point  $x \in F_m \cap \mathcal{P}(x_i)$  returns to  $\mathcal{P}(x_i)$ , and hence to  $B(x_i, \rho/2)$ . Recall that to  $x$  there is associated a backward branch  $\widehat{x}$  with  $\pi\widehat{x} = x$  and by (19) we have  $\widehat{f}^m(\widehat{x}) \in \widehat{\Lambda}_1$ . Given  $x \in F_m \cap \mathcal{P}(x_i)$ , let

$$U_x \stackrel{\text{def}}{=} f_x^{-m}(B(x_i, \rho/2)).$$

Notice that  $f^m(x) \in B(x_i, \rho/2) \subset B(f^m(x), \rho)$ . Hence the uniform hyperbolic estimates in (14) and (15) apply, and we can conclude that for every  $y \in B(x_i, \rho/2)$  and each  $0 \leq k \leq m-1$  we have  $\|df_x^{-k}(y)\| \leq e^{-k(\chi-2\varepsilon)}$  and

$$\text{diam } U_x = \text{diam } f_x^{-m}(B(x_i, \rho/2)) < e^{-k(\chi-2\varepsilon)}\rho.$$

This implies together with (20)

$$\text{diam } U_x \leq e^{-m(\chi-2\varepsilon)}\rho \leq e^{-n(\chi-2\varepsilon)}\rho \leq e^{-n_3(\chi-2\varepsilon)}\rho \leq \frac{1}{4}\rho \quad (23)$$

and hence  $U_x \subset B(x, \frac{1}{4}\rho)$  and

$$\overline{U_x} = \overline{f_x^{-m}(B(x_i, \rho/2))} \subset B(x_i, \rho).$$

For every two distinct points  $x, y$  in the  $(n, \rho/2)$ -separated set  $F_m$  in particular  $d(x, y) \geq \rho/2$  and hence (23) implies  $\overline{U_x} \cap \overline{U_y} = \emptyset$ . We observe that

$$R_{\varepsilon,0} \stackrel{\text{def}}{=} \overline{B(x_i, \rho/2)}, \quad R_{\varepsilon,\ell+1} \stackrel{\text{def}}{=} \bigcup_{x \in F_m} f_x^{-m}(R_{\varepsilon,\ell}) \quad \text{for } \ell \geq 0$$

form a family of nested non-empty compact sets, and hence define a non-empty compact set

$$R_\varepsilon \stackrel{\text{def}}{=} \bigcap_{\ell \geq 1} R_{\varepsilon,\ell} \quad (24)$$

that is, by construction,  $f^m$ -invariant. By (14) we have for every  $y \in R_\varepsilon$

$$\|df^m(y)(v)\| \geq e^{m(\chi-2\varepsilon)}\|v\|. \quad (25)$$

Hence  $f^m|_{R_\varepsilon}$  is uniformly expanding. Moreover,  $f^m|_{R_\varepsilon}$  is topologically conjugate to the one-sided full shift on an alphabet with  $\text{card } F_m$  symbols. This implies  $h_{\text{top}}(f^m|_{R_\varepsilon}) = \log \text{card } F_m$ , which implies for the set  $Q_\varepsilon \stackrel{\text{def}}{=} R_\varepsilon \cup f(R_\varepsilon) \cup \dots \cup f^{m-1}(R_\varepsilon)$

$$h_{\text{top}}(f|_{Q_\varepsilon}) = \frac{1}{m} \log \text{card } F_m.$$

Using  $(1+\varepsilon)n > m \geq n$ , (21), (22), and (17), we now obtain

$$\begin{aligned} h_{\text{top}}(f|_{Q_\varepsilon}) &\geq \frac{1}{m} \log \frac{1}{j} + \frac{n}{m} (h_\mu(f) - 2\varepsilon) \\ &> -\varepsilon + \frac{1}{1+\varepsilon} (h_\mu(f) - 2\varepsilon) \\ &\geq h_\mu(f) - 3\varepsilon. \end{aligned}$$

This proves property (a). From the construction we obtain that  $f|_{Q_\varepsilon}$  is a uniformly expanding repeller that satisfies (d). Because of Lemma 6 b) and (18) for every  $i = 1, \dots, K$  and every  $x \in Q_\varepsilon$  we have

$$\lim_{n \rightarrow \infty} \left| \frac{1}{n} (\varphi_i(x) + \varphi_i(f(x)) + \dots + \varphi_i(f^{n-1}(x))) - \int \varphi_i d\mu \right| \leq 2\varepsilon \quad (26)$$



and hence property (c). In particular, every  $f^m$ -invariant ergodic measure  $\nu$  supported on  $R_\varepsilon$  satisfies

$$\left| \frac{1}{m} \int S_m \varphi_i d\nu - \int \varphi_i d\mu \right| \leq 2\varepsilon. \quad (27)$$

Let  $\nu$  be an  $f^m$ -invariant ergodic measure supported on  $R_\varepsilon$  that has maximal entropy  $h_{\text{top}}(f^m|_{R_\varepsilon}) = h_\nu(f^m)$ . The variational principle for pressure and (27) together imply for every  $i = 1, \dots, K$

$$\frac{1}{m} P_{\text{top}}(f^m|_{R_\varepsilon}, S_m \varphi_i) \geq h_\nu(f) + \frac{1}{m} \int S_m \varphi_i d\nu - 3\varepsilon \geq h_\mu(f) + \int \varphi_i d\mu - 5\varepsilon.$$

Since, by [27, Theorem 9.8] we have

$$mP_{\text{top}}(f|_{Q_\varepsilon}, \varphi_i) = P_{\text{top}}(f^m|_{Q_\varepsilon}, S_m \varphi_i) \geq P_{\text{top}}(f^m|_{R_\varepsilon}, S_m \varphi_i),$$

we also have shown property (b). This finishes the proof of the theorem.

## REFERENCES

- [1] J. Alves, *SRB measures for non-hyperbolic systems with multidimensional expansion*, Ann. Scient. Éc. Norm. Sup. **33** (2000), 1–32.
- [2] J. Alves and V. Araujo, *Hyperbolic times: frequency versus integrability*, Ergodic Theory Dynam. Systems **24** (2004), 329–346.
- [3] J. Alves, Ch. Bonatti, and M. Viana, *SRB measures for partially hyperbolic systems whose central direction is mostly expanding*, Invent. Math. **140** (2000), 351–398.
- [4] L. Barreira and Ya. Pesin, *Lyapunov Exponents and Smooth Ergodic Theory*, Univ. Lecture Ser. **23**, Amer. Math. Soc., 2002.
- [5] J. Buzzzi, *On entropy-expanding maps*, preprint.
- [6] J. Buzzzi, O. Sester, and M. Tsujii, *Weakly expanding skew-products of quadratic maps*, Ergodic Theory Dynam. Systems **23** (2003), 1401–1414.
- [7] Y. M. Chung, *Shadowing properties of non-invertible maps with hyperbolic measures*, Tokyo J. Math. **22** (1999), 145–166.
- [8] M. Benedicks and M. Misiurewicz, *Absolutely continuous invariant measures for maps with flat tops*, Publ. Math. Inst. Hautes Étud. Sci. **69** (1989), 203–213.
- [9] A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Publ. Math. Inst. Hautes Étud. Sci. **51** (1980), 137–173.
- [10] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, *Encyclopedia of Mathematics and Its Applications* **54**, Cambridge University Press, 1995.
- [11] A. Katok and J.-M. Strelcyn, *Invariant Manifolds, Entropy and Billiards; Smooth Maps with Singularities*, *With the collab. of F. Ledrappier and F. Przytycki*, Lect. Notes Math. **1222**, Springer, 1986.
- [12] F. Ledrappier, *Some properties of absolutely continuous invariant measures on an interval*, Ergodic Theory Dynam. Systems **1** (1981), 77–93.
- [13] P.-D. Liu, *Pesin’s entropy formula for endomorphisms*, Nagoya Math. J. **150** (1998), 197–209.
- [14] S. Luzzatto and W. Tucker, *Non-uniformly expanding dynamics in maps with singularities and criticalities*, Publ. Math. Inst. Hautes Étud. Sci. **89** (1999), 179–226.
- [15] L. Mendoza, *Ergodic attractors for diffeomorphisms of surfaces*, J. London Math. Soc. (2) **37** (1988), 362–374.
- [16] R. Mañé, *Lyapunov exponents and stable manifolds for compact transformations*, Geometric dynamics, Proc. int. Symp., Rio de Janeiro/Brasil 1981, Lect. Notes Math. **1007**, 522–577 (1983).

- [17] M. Misiurewicz and W. Szlenk, *Entropy of piecewise monotone mappings*, Studia Math. **67** (1980), 45–63.
- [18] S. Newhouse, *Entropy and volume*, Ergodic Th. Dynam. Systems **8** (1988), 283–299.
- [19] T. Persson and J. Schmeling, *Dyadic diophantine approximation and Katok’s horseshoe approximation*, Acta Arith. **132** (2008), 205–230.
- [20] F. Przytycki, *Expanding repellers in limit sets for iterations of holomorphic functions*, Fund. Math. **186** (2005), 85–96.
- [21] F. Przytycki and M. Urbański, *Conformal Fractals: Ergodic Theory Methods*, London Mathematical Society Lecture Note Series 371, Cambridge University Press, 2010.
- [22] D. Ruelle, *Ergodic theory of differentiable dynamical systems*, Publ. Math. Inst. Hautes Études Sci. **50** (1979), 27–58.
- [23] F. Sánchez-Salas, *Ergodic attractors as limits of hyperbolic horseshoes*, Ergodic Theory Dyn. Syst. **22** (2002), 571–589.
- [24] H. Thunberg, *Positive exponent in families with flat critical points*, Ergodic Theory Dynam. Systems **19** (1999), 767–807.
- [25] V. A. Rokhlin, *Exact endomorphisms of a Lebesgue space*, Am. Math. Soc., Transl., II. Ser. **39** (1964), 1–36.
- [26] M. Viana, *Multidimensional non-hyperbolic attractors*, Inst. Hautes Etudes Sci. Publ. Math. **85** (1997), 63–96.
- [27] P. Walters, *An Introduction to Ergodic Theory*, Springer, 1982.
- [28] R. Zweimüller, *S-unimodal Misiurewicz maps with flat critical points*, Fund. Math. **181** (2004), 1–25.

INSTITUTO DE MATEMÁTICA, UFRJ, CIDADE UNIVERSITÁRIA - ILHA DO FUNDÃO,  
RIO DE JANEIRO 21945-909, BRAZIL

*E-mail address:* `katrin.gelfert@googlemail.com`